PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
G01N 33/68, 33/577, C12Q 1/68

A1

(11) International Publication Number: WO 00/26675

(43) International Publication Date: 11 May 2000 (11.05.00)

(21) International Application Number: PCT/CA99/01038

(22) International Filing Date: 3 November 1999 (03.11.99)

(30) Priority Data:
2,252,727 3 November 1998 (03.11.98) CA

(71) Applicant *(for all designated States except US)*: MCGILL UNIVERSITY [CA/CA]; 3550 University Street, Montreal, Quebec H3A 2A7 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ROULEAU, Guy [CA/CA]; 4850 Côte St-Luc, Apt.#7, Montreal, Quebec, H3W 2H2 (CA). JOOBER, Ridha [TN/CA]; 5282 Connaught, Montreal, Quebec H4V 1X6 (CA).

(74) Agents: DUBUC, Jean, H. et al.; Goudreau Gage Dubuc & Martineau Walker, The Stock Exchange Tower, Suite 3400, 800 Place Victoria, P.O. Box 242, Montreal, Quebec H4Z 1E9 (CA).

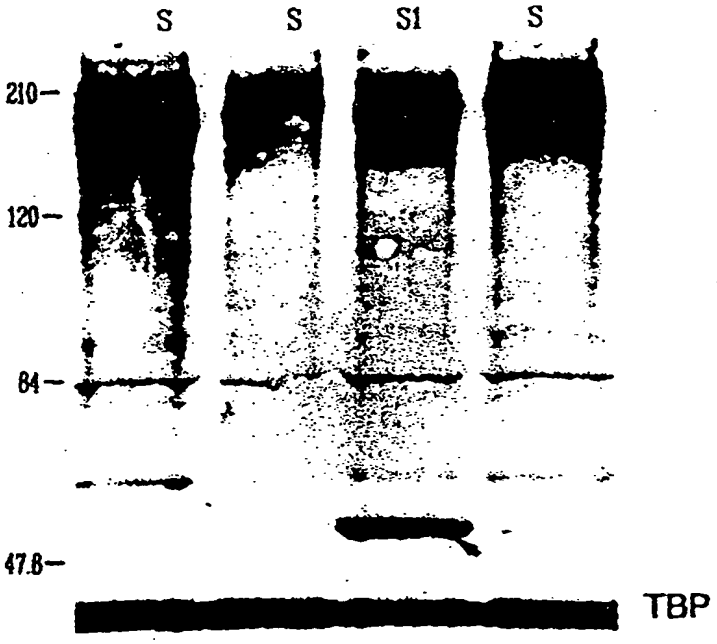
(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: POLYGLUTAMINE-CONTAINING PROTEINS IN NEUROPSYCHIATRIC DISORDERS

(57) Abstract

The present invention relates to diseases caused by CAG repeat expansions. More particularly, the present invention relates to neuropsychiatric diseases caused by CAG repeat expansions and even more particularly to schizophrenia and major depression. In particular, the present invention relates to the identification of at least one protein containing polyglutamine arrays, which acts as a marker for the pathogenesis of a neuropsychiatric disorder and more particularly for schizophrenia and major depression. The present invention further relates to diagnosis and treatment of neuropsychiatric disorders linked to polyglutamine tracts-containing proteins. The present invention further relates to diagnosis and treatment methods of neuropsychiatric disorders based on a targeting of the CAG/CTG repeat expansions encoding a polyglutamine tract. The present invention further provides methods of detecting nucleic acids encoding the polyglutamine tracts in a sample; kits containing nucleic acid probes or ligands of the polyglutamine tract, bio-assays using the nucleic acid sequences, and protein sequences of the present invention or ligands specific thereto to diagnose, assess, or prognose a mammal afflicted with or predisposed to develop a neuropsychiatric disorder linked to the presence of polyglutamine-containing proteins.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE OF THE INVENTION

POLYGLUTAMINE-CONTAINING PROTEINS IN NEUROPSYCHIATRIC DISORDERS

FIELD OF THE INVENTION

The present invention relates to diseases caused by CAG repeat expansions. More particularly, the present invention relates to neuropsychiatric diseases caused by CAG repeat expansions and even more particularly to schizophrenia and major depression. In particular, the present invention relates to the identification of at least one protein containing polyglutamine arrays, which acts as a marker for the pathogenesis of a neuropsychiatric disorder and more particularly for schizophrenia and major depression. The present invention further relates to diagnosis and treatment of neuropsychiatric disorders linked to polyglutamine tracts-containing proteins. The present invention further relates to diagnosis and treatment methods of neuropsychiatric disorders based on a targeting of the CAG/CTG repeat expansions encoding a polyglutamine tract. The present invention further provides methods of detecting nucleic acids encoding the polyglutamine tracts in a sample; kits containing nucleic acid probes or ligands of the polyglutamine tract, bio-assays using the nucleic acid sequences, and protein sequences of the present invention or ligands specific thereto to diagnose, assess, or prognose a mammal afflicted with or predisposed to develop a neuropsychiatric disorder linked to the presence of polyglutamine-containing proteins.

BACKGROUND OF THE INVENTION

Schizophrenia is a major psychiatric disorder that affects up to 1% of the general population. Family [1], twin [2] and adoption [3] studies indicate that genes play a significant role in its etiology. However, no genes or loci increasing susceptibility to schizophrenia have been convincingly identified.

Genetic anticipation (i.e. earlier age at onset and increased severity of the disease in successive generations) has been reported in schizophrenia [4-8]. Genetic anticipation is observed in a group of neurodegenerative diseases caused by CAG repeat expansions. The severity of these diseases correlates with the size of the trinucleotide expansions, which are unstable during gametogenesis. It was therefore hypothesized that genetic anticipation observed in some families with schizophrenia may be due to unstable CAG repeat expansions [9,10]. Consistent with this hypothesis, several studies using the repeat expansion detection method reported a higher average maximum size of CAG/CTG in schizophrenic patients compared to normal controls [11-13].

Although the exact role of expanded CAG repeats in the pathogenesis of neurodegenerative diseases is not clear, it is believed that their protein products are toxic to neural cells. Recently, it was shown that expanded polyglutamine-containing proteins (the amino acid glutamine is encoded by the codon CAG) tend to aggregate and form nuclear deposits which may represent one of the mechanisms leading to neural cell death [14].

If CAG repeat expansions are involved in schizophrenia, they may also act through expanded polyglutamine tracts. Such a correlation between the CAG repeats and polyglutamine stretches in protein from neuropsychiactric patients has yet to be provided.

There, thus remains a need to assess whether CAG/CTG repeats correlate with polyglutamine stretches in proteins obtained from CAG repeat-linked neuropsychiatric disorders. In addition, there remains a need to assess whether polyglutamine tracts containing proteins have a deleterious effects on cells. Further, there remains a need to assess whether polyglutamine containing proteins are valid markers for neuropsychiatric disorders. There also remains a need to provide means to treat neuropsychiatric diseases or conditions associated with CAG/CTG expansions.

The present invention seeks to meet these and other needs.

The present description refers to a number of documents, the content of which is herein incorporated by reference in their entirety.

SUMMARY OF THE INVENTION

The invention concerns the identification of polyglutamine containing proteins as a marker for CAG repeat-linked disorders. More particularly, the present invention concerns polyglutamine containing proteins as a marker for neuropsychiatric disorders, and more particularly for schizophrenia and major depression.

The invention further relates to CAG/CTG expansions and polyglutamine tracts as markers for neuropsychiatric disorders. In addition, the invention relates to CAG/CTG expansions and polyglutamine tracts as targets for diagnosing or treating neuropsychiatric disorders

The applicant was the first to demonstrate that polyglutamine containing proteins, encoded by CAG/CTG repeats and associated with a neuropsychiatric disease have deleterious effects on cells, as evidenced by the appearence of intra-nuclear inclusions in such diseased cells. The applicant is thus the first to have demonstrated that the polyglutamine tract containing proteins or proteins comprising an amino acid stretch encoded by a CAG/CTG repeat is a valid marker for a neuropsychiatric disease or condition, and especially for schizophrenia and major depression.

Before the present invention, it was unknown whether polyglutamine containing proteins could cause schizophrenia or major depression.

It shall also be understood, that since it has now been demonstrated that CAG repeats and polyglutamine tracts are causative agents of schizophrenia, the person of ordinary skill, will be able to adapt this teaching to other neuropsychiatric disorders.

While the diagnostic method of the instant invention is demonstrated with monoclonal antibody 1C2, other antibodies specific to

polyglutamine tracts could also be used. Such antibodies have been described for example in references 15 and 16. In addition, it is foreseeable that nucleic acid-based diagnostic methods (instead of protein-based, with the antibody) could be used to identify patients at risk of developping a neuropsychiatric disorder. It will be clear to the person of ordinary skill that the validation of a polyglutamine-containing protein in schizophrenia and major depression can be adapted to other neuropsychiatric disorders by a number of means.

In accordance with the present invention, there is thus provided a method of diagnosing the presence or predisposition to develop a neuropsychiatric disorder comprising: a) taking a sample from a patient; b) determining the presence of polyglutamine containing proteins in the sample; and c) diagnosing the presence or predisposition to develop a neuropsychiatric disorder, wherein a presence of polyglutamine containing proteins in the sample as compared to a sample from a patient without a neuropsychiatric disorder is indicative of the presence or predisposition to develop this neuropsychiatric disorder.

In accordance with the present invention, there is also provided a diagnostic kit comprising: a) a first container means containing a ligand specific to a polyglutamine array in a protein; and b) a second container means containing a conjugate comprising a binding partner of the ligand and a detectable label.

In addition, in accordance with the present invention, there is provided a method to identify modulators of the biological activity of polyglutamine-containing proteins associated with a neuropsychiatric disease. The method comprises: a) an incubation of cells from a patient harboring the polyglutamine-containing proteins with a molecule; and b) assessing the effect of the molecule on the biological activity of the polyglutamine containing proteins, wherein a difference in the biological activity in the presence of the molecule as compared to in its absence is indicative of a modulating activity of the molecule.

Furthermore, in accordance with the present invention, there is provided a method of identification of the cell type implicated in the development of a neuropsychiatric disorder comprising an identification of the cell type containing INI in a patient or animal suffering from a neuropsychiatric disease as compared to its absence in the cell type in a control patient or animal.

In accordance with the present invention, there is also provided a method of diagnosing the presence or predisposition to develop a neuropsychiatric disorder comprising: a) taking a sample from a patient; b) determining the presence of an amino acid stretch in a protein encoded by a CAG/CTG repeat in the sample; and c) diagnosing the presence or predisposition to develop a neuropsychiatric disorder, wherein a presence of the amino acid stretch encoded by the CAG/CTG repeat in the sample as compared to a sample from a patient without a neuropsychiatric disorder is indicative of the presence or predisposition to develop this neuropsychiatric disorder.

In order to provide a clear and consistent understanding of terms used in the present description, a number of definitions are provided hereinbelow.

GENERAL DEFINITIONS

Nucleotide sequences are presented herein by single strand, in the 5' to 3' direction, from left to right, using the one letter nucleotide symbols as commonly used in the art and in accordance with the recommendations of the IUPAC-IUB Biochemical Nomenclature Commission.

Unless defined otherwise, the scientific and technological terms and nomenclature used herein have the same meaning as commonly understood by a person of ordinary skill to which this invention pertains. Generally, the procedures for cell cultures, infection, molecular biology methods and the like are common methods used in the art. Such standard techniques can be found in reference manuals such as for example Sambrook et al. (1989,

Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratories) and Ausubel et al. (1994, Current Protocols in Molecular Biology, Wiley, New York).

The present description refers to a number of routinely used recombinant DNA (rDNA) technology terms. Nevertheless, definitions of selected examples of such rDNA terms are provided for clarity and consistency.

As used herein, "neurodegenerative disorders" (or diseases) relate to diseases in which brain cells degenerate, lose their function and eventually die. A non-limiting example of a neuro-degenerative disease is Huntington's disease or spinocerebellar ataxia type. "Neuropsychiatric disorders" or diseases refer to diseases or conditions relating to the psychiatric state of the affected individual. Non-limiting examples of neuropsychiatric disorders include schizophrenia, manic depression illness, autism, Tourette syndrom and restless leg syndrome.

As used herein, "nucleic acid molecule", refers to a polymer of nucleotides. Non-limiting examples thereof include DNA (i.e. genomic DNA, cDNA) and RNA molecules (i.e. mRNA). The nucleic acid molecule can be obtained by cloning techniques or synthesized. DNA can be double-stranded or single-stranded (coding strand or non-coding strand [antisense]).

The term "recombinant DNA" as known in the art refers to a DNA molecule resulting from the joining of DNA segments. This is often referred to as genetic engineering.

The term "DNA segment", is used herein, to refer to a DNA molecule comprising a linear stretch or sequence of nucleotides. This sequence when read in accordance with the genetic code, can encode a linear stretch or sequence of amino acids which can be referred to as a polypeptide, protein, protein fragment and the like.

The nucleic acid (i.e. DNA or RNA) for practicing the present invention may be obtained according to well known methods.

Oligonucleotide probes or primers of the present invention may be of any suitable length, depending on the particular assay format and the

particular needs and targeted genomes employed. In general, the oligonucleotide probes or primers are at least 12 nucleotides in length, preferably between 15 and 24 molecules, and they may be adapted to be especially suited to a chosen nucleic acid amplification system. As commonly known in the art, the oligonucleotide probes and primers can be designed by taking into consideration the melting point of hydrizidation thereof with its targeted sequence (see below and in Sambrook et al., 1989, Molecular Cloning - A Laboratory Manual, 2nd Edition, CSH Laboratories; Ausubel et al., 1989, in Current Protocols in Molecular Biology, John Wiley & Sons Inc., N.Y.).

The term "oligonucleotide" or "DNA" molecule or sequence refers to a molecule comprised of the deoxyribonucleotides adenine (A), guanine (G), thymine (T) and/or cytosine (C), in a double-stranded form, and comprises or includes a "regulatory element" according to the present invention, as the term is defined herein. The term "oligonucleotide" or "DNA" can be found in linear DNA molecules or fragments, viruses, plasmids, vectors, chromosomes or synthetically derived DNA. As used herein, particular double-stranded DNA sequences may be described according to the normal convention of giving only the sequence in the 5' to 3' direction.

"Nucleic acid hybridization" refers generally to the hybridization of two single-stranded nucleic acid molecules having complementary base sequences, which under appropriate conditions will form a thermodynamically favored double-stranded structure. Examples of hybridization conditions can be found in the two laboratory manuals referred above (Sambrook et al., 1989, *supra* and Ausubel et al., 1989, *supra*) and are commonly known in the art. In the case of a hybridization to a nitrocellulose filter, as for example in the well known Southern blotting procedure, a nitrocellulose filter can be incubated overnight at 65°C with a labeled probe in a solution containing 50% formamide, high salt (5 x SSC or 5 x SSPE), 5 x Denhardt's solution, 1% SDS, and 100 µg/ml denatured carrier DNA (i.e. salmon sperm DNA). The non-specifically binding probe can then be washed off the

filter by several washes in 0.2 x SSC/0.1% SDS at a temperature which is selected in view of the desired stringency: room temperature (low stringency), 42°C (moderate stringency) or 65°C (high stringency). The selected temperature is based on the melting temperature (Tm) of the DNA hybrid. Of course, RNA-DNA hybrids can also be formed and detected. In such cases, the conditions of hybridization and washing can be adapted according to well known methods by the person of ordinary skill. Stringent conditions will be preferably used (Sambrook et al.,1989, *supra*).

Probes of the invention can be utilized with naturally occurring sugar-phosphate backbones as well as modified backbones including phosphorothioates, dithionates, alkyl phosphonates and α-nucleotides and the like. Modified sugar-phosphate backbones are generally taught by Miller, 1988, Ann. Reports Med. Chem. 23:295 and Moran et al., 1987, Nucleic acid molecule. Acids Res., 14:5019. Probes of the invention can be constructed of either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), and preferably of DNA.

The types of detection methods in which probes can be used include Southern blots (DNA detection), dot or slot blots (DNA, RNA), and Northern blots (RNA detection). Although less preferred, labeled proteins could also be used to detect a particular nucleic acid sequence to which it binds. Other detection methods include kits containing probes on a dipstick setup and the like.

Although the present invention is not specifically dependent on the use of a label for the detection of a particular nucleic acid sequence, such a label might be beneficial, by increasing the sensitivity of the detection. Furthermore, it enables automation. Probes can be labeled according to numerous well known methods (Sambrook et al., 1989, supra). Non-limiting examples of labels include ^{3}H, ^{14}C, ^{32}P, and ^{35}S. Non-limiting examples of detectable markers include ligands, fluorophores, chemiluminescent agents, enzymes, and antibodies. Other detectable markers for use with probes, which can enable an increase in sensitivity of the method of the invention, include

biotin and radionucleotides. It will become evident to the person of ordinary skill that the choice of a particular label dictates the manner in which it is bound to the probe.

As commonly known, radioactive nucleotides can be incorporated into probes of the invention by several methods. Non-limiting examples thereof include kinasing the 5' ends of the probes using gamma ^{32}P ATP and polynucleotide kinase, using the Klenow fragment of Pol I of *E. coli* in the presence of radioactive dNTP (i.e. uniformly labeled DNA probe using random oligonucleotide primers in low-melt gels), using the SP6/T7 system to transcribe a DNA segment in the presence of one or more radioactive NTP, and the like.

As used herein, "oligonucleotides" or "oligos" define a molecule having two or more nucleotides (ribo or deoxyribonucleotides). The size of the oligo will be dictated by the particular situation and ultimately on the particular use thereof and adapted accordingly by the person of ordinary skill. An oligonucleotide can be synthetised chemically or derived by cloning according to well known methods.

As used herein, a "primer" defines an oligonucleotide which is capable of annealing to a target sequence, thereby creating a double stranded region which can serve as an initiation point for DNA synthesis under suitable conditions. In a particular embodiment one such primer will anneal to CAG/CTG expansions.

Amplification of a selected, or target, nucleic acid sequence may be carried out by a number of suitable methods. See generally Kwoh et al., 1990, Am. Biotechnol. Lab. 8:14-25. Numerous amplification techniques have been described and can be readily adapted to suit particular needs of a person of ordinary skill. Non-limiting examples of amplification techniques include polymerase chain reaction (PCR), ligase chain reaction (LCR), strand displacement amplification (SDA), transcription-based amplification, the Qβ replicase system and NASBA (Kwoh et al., 1989, Proc. Natl. Acad. Sci. USA 86,

1173-1177; Lizardi et al., 1988, BioTechnology 6:1197-1202; Malek et al., 1994, Methods Mol. Biol., 28:253-260; and Sambrook et al., 1989, *supra*). Preferably, amplification will be carried out using PCR.

Polymerase chain reaction (PCR) is carried out in accordance with known techniques. See, e.g., U.S. Pat. Nos. 4,683,195; 4,683,202; 4,800,159; and 4,965,188 (the disclosures of all three U.S. Patent are incorporated herein by reference). In general, PCR involves, a treatment of a nucleic acid sample (e.g., in the presence of a heat stable DNA polymerase) under hybridizing conditions, with one oligonucleotide primer for each strand of the specific sequence to be detected. An extension product of each primer which is synthesized is complementary to each of the two nucleic acid strands, with the primers sufficiently complementary to each strand of the specific sequence to hybridize therewith. The extension product synthesized from each primer can also serve as a template for further synthesis of extension products using the same primers. Following a sufficient number of rounds of synthesis of extension products, the sample is analysed to assess whether the sequence or sequences to be detected are present. Detection of the amplified sequence may be carried out by visualization following EtBr staining of the DNA following gel electrophores, or using a detectable label in accordance with known techniques, and the like. For a review on PCR techniques (see PCR Protocols, A Guide to Methods and Amplifications, Michael et al. Eds, Acad. Press, 1990).

Ligase chain reaction (LCR) is carried out in accordance with known techniques (Weiss, 1991, Science 254:1292). Adaptation of the protocol to meet the desired needs can be carried out by a person of ordinary skill. Strand displacement amplification (SDA) is also carried out in accordance with known techniques or adaptations thereof to meet the particular needs (Walker et al., 1992, Proc. Natl. Acad. Sci. USA 89:392-396; and ibid., 1992, Nucleic Acids Res. 20:1691-1696).

As used herein, the term "gene" is well known in the art and relates to a nucleic acid sequence defining a single protein or polypeptide. A

"structural gene" defines a DNA sequence which is transcribed into RNA and translated into a protein having a specific amino acid sequence thereby giving rise the a specific polypeptide or protein. It will be readily recognized by the person of ordinary skill, that the nucleic acid sequence of the present invention can be incorporated into anyone of numerous established kit formats which are well known in the art.

A "heterologous" (i.e. a heterologous gene) region of a DNA molecule is a subsegment segment of DNA within a larger segment that is not found in association therewith in nature. The term "heterologous" can be similarly used to define two polypeptidic segments not joined together in nature. Non-limiting examples of heterologous genes include reporter genes such as luciferase, chloramphenicol acetyl transferase, β-galactosidase, and the like which can be juxtaposed or joined to heterologous control regions or to heterologous polypeptides.

The term "vector" is commonly known in the art and defines a plasmid DNA, phage DNA, viral DNA and the like, which can serve as a DNA vehicle into which DNA of the present invention can be cloned. Numerous types of vectors exist and are well known in the art.

The term "expression" defines the process by which a gene is transcribed into mRNA (transcription), the mRNA is then being translated (translation) into one polypeptide (or protein) or more.

The terminology "expression vector" defines a vector or vehicle as described above but designed to enable the expression of an inserted sequence following transformation into a host. The cloned gene (inserted sequence) is usually placed under the control of control element sequences such as promoter sequences. The placing of a cloned gene under such control sequences is often refered to as being operably linked to control elements or sequences.

Operably linked sequences may also include two segments that are transcribed onto the same RNA transcript. Thus, two sequences, such

as a promoter and a "reporter sequence" are operably linked if transcription commencing in the promoter will produce an RNA transcript of the reporter sequence. In order to be "operably linked" it is not necessary that two sequences be immediately adjacent to one another.

Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host or both (shuttle vectors) and can additionally contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements, and/or translational initiation and termination sites.

In accordance with one embodiment of the present invention, an expression vector can be constructed to assess the deleterious effects of CAG repeats in cells or animals. For example, the deleterious effects of CAG repeats could be analyzed by assessing the formation of INI following expression of proteins comprising an amino acid stretch encoded by CAG repeats. Of course, through recombinant methods, different lengths and types of CAG repeats could be tested. In addition, expression vectors having been shown to trigger INI formation in cells or animals, could be used to test and identify agents which diminish or abrogate the CAG-repeat-dependent deleterious effects. Non-human transgenic animals expressing chosen proteins harboring a polypeptide portion encoded by a CAG repeat could also be prepared and used to screen compounds that diminish or abrogate the deleterious effects thereof.

Prokaryotic expressions are useful for the preparation of large quantities of the protein encoded by the DNA sequence of interest. This protein can be purified according to standard protocols that take advantage of the intrinsic properties thereof, such as size and charge (i.e. SDS gel electrophoresis, gel filtration, centrifugation, ion exchange chromatography...). In addition, the protein of interest can be purified via affinity chromatography using polyclonal or monoclonal antibodies. The purified protein can be used for therapeutic applications.

As used herein, the designation "functional derivative" denotes, in the context of a functional derivative of a sequence whether a nucleic acid or amino acid sequence, a molecule that retains a biological activity (either function or structural) that is substantially similar to that of the original sequence. This functional derivative or equivalent may be a natural derivative or may be prepared synthetically. Such derivatives include amino acid sequences having substitutions, deletions, or additions of one or more amino acids, provided that the biological activity of the protein is conserved. The same applies to derivatives of nucleic acid sequences which can have substitutions, deletions, or additions of one or more nucleotides, provided that the biological activity of the sequence is generally maintained. When relating to a protein sequence, the substituting amino acid as chemico-physical properties which are similar to that of the substituted amino acid. The similar chemico-physical properties include, similarities in charge, bulkiness, hydrophobicity, hydrophylicity and the like. The term "functional derivatives" is intended to include "fragments", "segments", "variants", "analogs" or "chemical derivatives" of the subject matter of the present invention.

Thus, the term "variant" refers herein to a protein or nucleic acid molecule which is substantially similar in structure and biological activity to the protein or nucleic acid of the present invention.

The functional derivatives of the present invention can be synthesized chemically or produced through recombinant DNA technology. All these methods are well known in the art.

As used herein, "chemical derivatives" is meant to cover additional chemical moieties not normally part of the subject matter of the invention. Such moieties could affect the physico-chemical characteristic of the derivative (i.e. solubility, absorption, half life and the like, decrease of toxicity). Such moieties are examplified in Remington's Pharmaceutical Sciences (1980). Methods of coupling these chemical-physical moieties to a polypeptide are well known in the art.

The term "allele" defines an alternative form of a gene which occupies a given locus on a chromosome.

As commonly known, a "mutation" is a detectable change in the genetic material which can be transmitted to a daughter cell. As well known, a mutation can be, for example, a detectable change in one or more deoxyribonucleotide. For example, nucleotides can be added, deleted, substituted for, inverted, or transposed to a new position. Spontaneous mutations and experimentally induced mutations exist. The result of a mutations of nucleic acid molecule is a mutant nucleic acid molecule. A mutant polypeptide can be encoded from this mutant nucleic acid molecule.

As used herein, the term "purified" refers to a molecule having been separated from a cellular component. Thus, for example, a "purified protein" has been purified to a level not found in nature. A "substantially pure" molecule is a molecule that is lacking in all other cellular components.

As used herein, the terms "molecule", "compound" or "ligand" are used interchangeably and broadly to refer to natural, synthetic or semi-synthetic molecules or compounds. The term "molecule" therefore denotes for example chemicals, macromolecules, cell or tissue extracts (from plants or animals) and the like. Non limiting examples of molecules include nucleic acid molecules, peptides, antibodies, carbohydrates and pharmaceutical agents. The agents can be selected and screened by a variety of means including random screening, rational selection and by rational design using for example protein or ligand modelling methods such as computer modelling. The terms "rationally selected" or "rationally designed" are meant to define compounds which have been chosen based on the configuration of the interaction domains of the present invention. As will be understood by the person of ordinary skill, macromolecules having non-naturally occurring modifications are also within the scope of the term "molecule". For example, peptidomimetics, well known in the pharmaceutical industry and generally referred to as peptide analogs can be generated by modelling as mentioned above. Similarly, in a preferred

embodiment, the polypeptides of the present invention are modified to enhance their stability. It should be understood that in most cases this modification should not alter the biological activity of the interaction domain. The molecules identified in accordance with the teachings of the present invention have a therapeutic value in diseases or conditions in which the physiology or homeastasis of the cell and/or tissue is compromised by the presence of a CAG/CGT expression or polyglutamine arrays resulting in a neuropsychiatric disease or condition. Alternatively, the molecules identified in accordance with the teachings of the present invention find utility in the development of more efficient molecules that could modulate the biological effets of these expansions or arrays.

As used herein, agonists and antagonists also include potentiators of known compounds with such agonist or antagonist properties. In one embodiment, modulators of the deleterious effects of proteins harboring an amino acid stretch encoded by a CAG/CTG repeat can be identified and selected by contacting the indicator cell with a compound, mixture thereof, or a library of molecules for a fixed period of time. In certain embodiments, the same protein lacking this amino acid stretch can be used as a positive control.

An indicator cell in accordance with the present invention can be used to identify antagonists. For example, the test molecule or molecules are incubated with the host cell in conjunction with one or more agonists held at a fixed concentration. An indication and relative strength of the antagonistic properties of the molecule(s) can be provided by comparing the number of INIs in the presence of the agonist, in the absence of test molecules vs in the presence thereof. Of course, the antagonistic effect of a molecule can also be determined in the absence of agonist, simply by comparing the number of INIs in the presence and absence of the test molecule(s).

As used herein the recitation "indicator cells" refers to cells that express a protein harboring an amino acid stretch encoded by a deleterious CAG repeat according to the present invention. As alluded to above, such

indicator cells can be used in a screening assay of the present invention. In certain embodiments, the indicator cells have been engineered so as to express a chosen derivative, fragment, homolog, or mutant of the protein harboring an amino acid stretch encoded by a deleterious CAG repeat according to the present invention. Such an indicator cell could be used to rapidly screen at high-throughput a vast array of test molecules.

A host cell or indicator cell has been "transfected" by exogenous or heterologous DNA (e.g. a DNA construct) when such DNA has been introduced inside the cell. The transfecting DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transfecting DNA may be maintained on a episomal element such as a plasmid. With respect to eukaryotic cells, a stably transfected cell is one in which the transfecting DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transfecting DNA. Transfection methods are well known in the art (Sambrook et al., 1989, *supra*; Ausubel et al., 1994 *supra*).

The present invention also provides antisense nucleic acid molecules which can be used for example to decrease or abrogate the expression of the nucleic acid sequences or proteins of the present invention. An antisense nucleic acid molecule according to the present invention refers to a molecule capable of forming a stable duplex or triplex with a portion of its targeted nucleic acid sequence (DNA or RNA). The use of antisense nucleic acid molecules and the design and modification of such molecules is well known in the art as described for example in WO 96/32966, WO 96/11266, WO 94/15646, WO 93/08845 and USP 5,593,974. Antisense nucleic acid molecules according to the present invention can be derived from the nucleic acid sequences and modified in accordance to well known methods. For example,

some antisense molecules can be designed to be more resistant to degradation to increase their affinity to their targeted sequence, to affect their transport to chosen cell types or cell compartments, and/or to enhance their lipid solubility bu using nucleotide analogs and/or substituting chosen chemical fragments thereof, as commonly known in the art.

In one particular embodiment, a polyglutamine stretch may be used to identify factors which interact therewith and are implicated in cell damage which leads to the neuropsychiatric disorder. In one particular embodiment, the polyglutamine stretch may be fused to a heterogogous polypeptide to yield a fusion protein. Non limiting examples of such fusion proteins include a hemaglutinin fusions and Gluthione-S-transferase (GST) fusions and Maltose binding protein (MBP) fusions. In certain embodiments, it might be beneficial to introduce a protease cleavage site between the two polypeptide sequences which have been fused. Such protease cleavage sites between two heterologously fused polypeptides are well known in the art.

In certain embodiments, it might also be beneficial to fuse the interaction domains of the present invention to signal peptide sequences enabling a secretion of the fusion protein from the host cell. Signal peptides from diverse organisms are well known in the art. Bacterial OmpA and yeast Suc2 are two non limiting examples of proteins containing signal sequences. In certain embodiments, it might also be beneficial to introduce a linker (commonly known) between the interaction domain and the heterologous polypeptide portion. Such fusion protein find utility in the assays of the present invention as well as for purification purposes, detection purposes and the like.

For certainty, the sequences and polypeptides useful to practice the invention include without being limited thereto mutants, homologs, subtypes, alleles and the like. It will be clear to the person of ordinary skill that whether the sequences domain of the present invention, variant, derivative, or fragment thereof retain their function (i.e.in triggering deleterious effects in cells and in being internalized in inclusion bodies) can be readily determined by using

the teachings and assays of the present invention and the general teachings of the art.

In general, techniques for preparing antibodies (including monoclonal antibodies and hybridomas) and for detecting antigens using antibodies are well known in the art (Campbell, 1984, In "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology", Elsevier Science Publisher, Amsterdam, The Netherlands) and in Harlow et al., 1988 (in: Antibody- A Laboratory Manual, CSH Laboratories). The present invention also provides polyclonal, monoclonal antibodies, or humanized versions thereof, chimeric antibodies and the like which recognize the polyglutamine arrays and/or the amino acid stretches of proteins encoded by the CAG/CTG repeats of the present invention and could inhibit or neutralize their deleterious effects in cells.

From the specification and appended claims, the term therapeutic agent should be taken in a broad sense so as to also include a combination of at least two such therapeutic agents. Further, the DNA segments or proteins according to the present invention can be introduced into individuals in a number of ways well known in the art. For administration to humans, the prescribing medical professional will ultimately determine the appropriate form and dosage for a given patient, and this can be expected to vary according to the chosen therapeutic regimen (i.e. DNA construct, protein, cells), the response and condition of the patient as well as the severity of the disease.

Composition within the scope of the present invention should contain the active agent (i.e. fusion protein, nucleic acid, antibody, and molecule) in an amount effective to achieve the desired therapeutic effect while avoiding adverse side effects. Typically, the therapeutic agent in accordance with the present invention can be administered to mammals (i.e. humans) in doses ranging from 0.005 to 1 mg per kg of body weight per day of the mammal which is treated. Pharmaceutically acceptable preparations and salts of the

active agent are within the scope of the present invention and are well known in the art (Remington's Pharmaceutical Science, 16th Ed., Mack Ed.). For the administration of polypeptides, antagonists, agonists and the like, the amount administered should be chosen so as to avoid adverse side effects. The dosage will be adapted by the clinician in accordance with conventional factors such as the extent of the disease and different parameters from the patient. Typically, 0.001 to 50 mg/kg/day will be administered to the mammal.

The present invention also relates to a kit for diagnosing a neuropsychiatric disease or condition or a predisposition to contracting same comprising a nucleic acid, a protein or a ligand in accordance with the present invention. For example, a compartmentalized kit in accordance with the present invention includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample (DNA or cells), a container which contains the primers used in the assay, containers which contain enzymes, containers which contain wash reagents, and containers which contain the reagents used to detect the extension products.

The present invention also provides the means to assess the time and place where the mutant proteins (polyglutamine tract containing proteins) of the present invention can be identified. It also provides the means to assess whether a threshold quantity thereof is necessary to trigger a neuropsychiatric disease. The identification of the time, place and level of expression of the mutant proteins of the present invention in brain will allow the identification of key cells and pathways in brain development which when perturbed or disrupted lead to a neuropsychiatric disorder. Since the brain cells

implicated in the development of neuropsychiatric diseases have yet to be identified. The instant invention opens the way to the identification of such cells.

BRIEF DESCRIPTION OF THE DRAWINGS

Having thus generally described the invention, reference will now be made to the accompanying drawings, showing by way of illustration a preferred embodiment thereof, and in which:

Figure 1 shows western blot immunoreactivity in a series of lymphoblastoid total protein extracts from lines established from schizophrenic patients (S), relatives of one schizophrenic patient (R) and healthy volunteers (C). All samples contained the same volume of cell extract. TBP: TATA binding protein. The Blot was exposed to the 1C2 antibody (50ml, 1/5000) overnight at 4°C. Positions of molecular size marker are indicated on the left of the figure. Arrowhead indicates the novel bands observed in the two unrelated schizophrenic patients (S1 and S4).

Figure 2 shows a western blot immunoreactivity in lymphoblastoid total protein extracts from lines established from 4 schizophrenic patients. All samples contained approximately 5 micrograms of proteins. TBP: TATA binding protein. Arrowhead shows the novel band observed in S1. The Blot was exposed to the 1C2 antibody (10ml, 1/2000) for 1 hour at room temperature. Positions of molecular size marker are indicated on the left of the figure.

Figure 3 shows the immunocytochemical detection of intranuclear inclusions (INI) in lymphoblastoid cells (LCL). Immuno-cytochemistry of control LCL versus schizophrenia LCL: absence of INI in control LCL probed with anti-ubiquitin (**a**); and detection of INI in schizophrenia LCL with anti-ubiquitin (**b** for S1 and **c** for S4). For all panels, the magnification before publication is 400x (left) and 1000x (right).

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawing which is exemplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Genetic anticipation, manifested by increased severity and earlier age at onset of the disease over successive generations, is reported in schizophrenia. The molecular basis of anticipation in several neurodegenerative diseases is unstable coding CAG repeat expansions. Anticipation was reported in schizophrenia. Recently, studies suggested that enlarged CAG/CTG repeats are over represented in schizophrenic patients compared to normal controls. Together, these observations suggest that unstable CAG repeats may play a role in the etiology of schizophrenia. In order to assess the derival relevance of such CAG repeats in patients, the presence of polyglutamine expanded tracts, encoded by CAG repeats, in total protein extracts derived from lymphoblastoid cell lines of schizophrenic patients was analysed. Proteins from schizophrenic patients (n = 57) and normal controls (n = 73) were separated by means of SDS-polyacrylamide gel electrophoresis, wet-blotted onto nitrocellulose membrane and probed with a monoclonal antibody (mab 1C2) recognizing expanded polyglutamine arrays. Three abnormal bands corresponding to protein(s) of molecular weight of approximately 50 KDa were identified in 2 unrelated schizophrenic patients and in a sibling of one of these patients who suffered a major depression. None of the normal controls tested positive for this abnormal band. These results suggest that expanded polyglutamine containing proteins, though rare, may play a role in the pathogenesis of schizophrenia.

Having identified polyglutamine-containing proteins in schizophrenic patients, it was then assessed whether the presence thereof could

be associated with deleterious effects in the form of accumulation of intranuclear inclusions (INI). Lymphoblasts from the above-mentioned patients were analyzed for the presence of INI. Surprisingly, INI were identified in these patients, but not the controls. The present invention thus provides a correlation between the presence of polyglutamine containing proteins and a neuropsychiatric disorder.

The present invention is illustrated in further detail by the following non-limiting examples.

EXAMPLE 1

POLYGLUTAMINE ARRAYS CONTAINING PROTEINS ARE ASSOCIATED WITH DELETERIOUS EFFECTS IN CELLS OF SCHIZOPHRENIA PATIENTS.

To formally test if CAG repeat expansions are involved in schizophrenia, a monoclonal antibody (mab1C2) that specifically recognizes polyglutamine tracts was used to identify polyglutamine-containing proteins in total protein extracts derived from lymphoblastoid cell lines of schizophrenic patients and normal controls. The 1C2 antibody, initially raised against a peptide that encompasses the polyglutamine stretch of the TATA binding protein (TBP) [15], was subsequently found to specifically recognize large (expanded [16] and non expanded [17]) polyglutamine arrays expressed in other proteins. The effect of these polyglutamine arrays containing proteins on cells was then analyzed.

METHODS

SUBJECT

The patients (N = 57) involved in this study are part of a currently ongoing pharmaco-genetic project. In an attempt to reduce genetic heterogeneity, schizophrenic patients were categorized on the basis of *a priori* defined criteria of severity of the disease and responsiveness to neuroleptic medication. Of course, the person skilled in the art will realize that other

categorization of the patients could be carried out. Neuroleptic-non-responsive patients (N = 28) were recruited according to the following criteria: (1) Chronic schizophrenia, according to the Diagnostic and Statistical Manual of Mental Diseases (DSM-IV [18]), (2) Continuous psychotic symptoms with no significant remission within the past 2 years, (3) At least 3 periods of treatment with typical neuroleptics, from at least two distinct families of drugs, at clinically sufficient dosage (equal to or greater than 750 mg, while treated in monotherapy, or 1000 mg chlorpromazine equivalent, while treated with two or more neuroleptics combined), for a continuous period of at least 6 weeks at a time, with no significant relief of symptoms in the preceding 5 years, and; (4) Unable to function without supervision in all or nearly all domains of social and vocational activities.

Criteria for neuroleptic-responsive patients (N = 29) were: (1) schizophrenia, according to DSM-IV, (2) at least one admission to a psychiatric care facility because of acute psychotic episode, (3) during all hospitalizations, full or partial remission in response to treatment with typical neuroleptics, at recommended dosage, within six-eight weeks of continuous treatment; the term *remission* is defined herein as a rapid reduction of schizophrenic symptoms with limited residual symptoms, (4) able to function with only occasional supervision in all or nearly all domains of social and vocational activities, (5) no admissions to hospitals because of psychotic exacerbation, if compliant to treatment and treated continuously with typical neuroleptics, and; (6) at least one psychotic relapse when neuroleptic medication is reduced or interrupted. Exclusion criteria for schizophrenic patients were brain trauma, any neurological condition and drug or alcohol abuse in the last two years.

Neuroleptic non-responsive patients were selected from a list of schizophrenic patients identified as candidates for treatment or treated with atypical neuroleptics because of treatment resistance. Three institutions provided these non-responsive patients: Douglas Hospital, Clinique Jeunes

Adultes of L.H. Lafontaine Hospital and the Schizophrenia Clinic of the Royal Ottawa Hospital. Responsive patients were selected from a list of all patients who were considered very good responders to neuroleptics by their treating physicians and/or nurse and who were followed in the out-patient clinics attached to the Douglas and L.H. Lafontaine hospitals.

Patients were directly interviewed using the Diagnostic Interview for Genetic Studies [19] (DIGS) and their medical records were comprehensively reviewed by a research psychiatrist. Complementary information from the treating physician and nurses in charge was obtained. Diagnosis was based on the concordance of two psychiatrists using DSM criteria and all the available data. Diagnoses in relatives were based on direct interview using the Diagnostics Interview DIGS and review of the medical files.

A group of 34 normal controls was recruited through local advertisement and among healthy married-in individuals from an ongoing linkage study. All controls have been screened for DSM-IV axis I disorders. In addition, 39 psychiatrically unscreened control subjects were tested. All patients and controls were Caucasians with western and central European descent.

After complete description of the study to the subjects, written informed consent was obtained.

Protein extraction and western blotting

Whole-cell protein extracts were isolated from lymphoblastoid cell lines transformed by Epstein-Barr Virus using standard techniques. The protein extraction was performed on ice and the PMSF protease inhibitor was added (1mM final concentration) to the protein homogenate. Fifty micrograms of proteins from each sample were separated using an 8% SDS-PAGE gel. Gels were blotted onto nitrocellulose transfer and immobilization membrane (Schleicher & Schuell, Keene, NH, USA) using BIO RAD Trans-Blotting electrophoretic transfer cell apparatus following manufacturer recommendations

(Biorad, Hercules, CA, USA). The quality of protein transfer and the amount of proteins was checked using the Ponseau red staining method. Blots were probed over night at 4°C with the primary antibody (mab1C2; 1:5000) then detected using the horseradish peroxidase-conjugated secondary antibodies (1:4000) (Jackson Immino research laboratories, baltimore USA) and enhanced chemiluminescence (ECL protocol, Amersham, Little Chalfont, Buckinghamshire, England).

Immunocytochemistry

Schizophrenia and control lymphoblastoid cells were harvested and a total of $50x10^4$ cells from each cell line were plated onto poly-D-lysine coated slides. Immunodetection was performed using anti-ubiquitin (1:300) antibody. Biotinylated secondary antibodies were used at a 1:500 dilution and an amplification step was performed using the ABC kit (VECTOR). Colorimetric detection was carried out using the VIP kit (VECTOR).

RESULTS

Many immunoreactive bands were observed in all schizophrenic patients and healthy volunteers indicating that there are proteins, as yet uncharacterized, detected by the 1C2 monoclonal antibody. The strongest signal, corresponding to the TATA binding protein (with a size <48 KD), was present in all subjects and may be considered as an internal positive control for each subject.

Two strongly immunoreactive bands were detected in two schizophrenic subjects (figure 1: respectively lanes S1 and S4) but not in the 73 controls. Although both of them are of approximate molecular weight of 50 KDa, the band observed in S4 seemed to be slightly smaller than the one observed in S1. To ensure the reproducibility of these bands, proteins were extracted from at least 2 different cell line aliquots established from the 2 patients, and western blots were repeated using different experimental stringency conditions. These bands were consistently identified under several experimental conditions.

The immunoreactivity of the 50 KDa band detected in S1 was equivalent to that of the TBP band when the blot was incubated at room temperature for 1 hour with the primary antibody at a 1:2000 concentration (see figure 2). In order to exclude the presence of immunoglobulins in these samples (S1 and S4) that would react with the secondary antibody, the immunodetection using only the secondary antibody was repeated. No abnormal bands were detected. Both patients were diagnosed with paranoid schizophrenia and were good responders to neuroleptic medication. S4 had no family history of mental disorders and his age at onset was 25 years (average of the group 21 ± 5 years). S1 had an earlier age at onset (19 years) and had 2 siblings (out of 8) who were diagnosed respectively with a major depressive episode and a postpartum depression. A novel slightly smaller band as that observed in the schizophrenic proband was detected in the sibling diagnosed with major depression (Figure 1 lane R2); no novel protein was detected in the sibling with postpartum depression. Three other non-affected siblings were tested and did not display any additional bands. No parents were available for testing (S1 parents were dead and S4 parents declined to participate).

Three other schizophrenic patients and one of the siblings of S1 (figure 1 lanes S2, S5, S6 and R2 respectively) showed weakly immunoreactive bands of smaller molecular weight (approximately 45-48 KD). However, fainter immunoreactive bands in the same size range were observed in normal volunteers (Figure 1 lane C2, C3, C6 and C7). The nature of these bands remains unclear.

Detection of intranuclear inclusions in schizophrenia patients with the novel 1C2 positive protein band

Coding CAG triplet repeat expansions cause several neurodegenerative disorders, including Machado-Joseph disease (MJD)[14, 23-26]. Intranuclear filamentous inclusions (INI) observed in MJD, as well as in other CAG repeat disorders, have lead to a nuclear toxicity model[14, 27-30].

The toxic waste management model in neuron posits that polyglutamine-containing proteins cause neurodegeneration because mutant forms of these proteins tend to aggregate in ubiquitin-positive inclusions in the nucleus Cohen 1998, (Nature Genetics 19:109; editorial box); Cummings et al. 1998, (Nature Genetics 19:148-154) have reported that ataxin-1 aggregates also stain positive for proteasome components and the HDJ-2/HSDJ chaperones. This raises the possibility that the cell may actively attempt to degrade the polyglutamine-containing ataxin-1 without a complete degradation thereof. The cellular waste-removal machinery would be monopolized such that degradation and removal of other proteins is compromized, thereby contributing and ultimately leading to cellular degeneration.

To test if the protein accumulates in the nuclei we performed immunocytochemistry on lymphoblasts from four controls and the two schizophrenia patients displaying a novel band with 1C2 were performed. Figure 3 depicts that ubiquitin positive INI are observed in schizophrenia cell lines and not in controls. These results have been replicated in three separate experiments.

Cummings et al., 1998 *(supra)* also show that overexpression of chaperone proteins expressing mutant ataxin-1 (polyglutamine containing ataxin-1) decreases the frequency of aggregate formation. Having now shown that polyglutamine proteins are also included into INI in a neuropsychiatric disorder (i.e. schizophrenia), opens the way to assessing whether overexpression of a chaperone protein also reduces aggregates in the diseased cells. Such an overexpression on the disease itself should also be tested with an animal model.

DISCUSSION

Two recent studies did not identify polyglutamine expansions in schizophrenic patients [20,21]. The first study by Schurhoff et al. [20] included only three schizophrenic patients, none of which had family history of schizophrenia

or related spectrum disorders. The second study by Jones et al. [21] included 18 schizophrenic patients where repeat expansion detection method showed CAG/CTG expanded genomic sequences. No expanded polyglutamine containing proteins were detected. However, in contrast with the present results, the only signal detected on the western blots was the TBP protein band, suggesting methodological differences with the present study. Indeed, as shown herein, several constant bands were observed. It is therefore possible that the results obtained by Jones et al. were due to different experimental conditions where only the TBP, the primary epitope of the 1C2 antibody, could be detected (such as 30 second exposure to the autoradiographic film in Jones et al. experiments compared to 1 minute in our experiments). Using very stringent conditions will increase the specificity of the experiment allowing only detection of very expanded polyglutamine tracts (or the TBP). However, these same stringent conditions will result in a low sensitivity preventing the detection of weaker antigens present in diseases with known CAG repeat expansions (Lopes-Cendes, unpublished data).

Although the 1C2 antibody detects preferentially expanded polyglutamine tracts, smaller polyglutamine polymers are also detected by this antibody [17]. It is therefore difficult to be certain that the additional protein bands observed in the present study correspond to expanded polyglutamine tracts in the same size range as those causing neurodegenerative diseases. However, polyglutamine peptides with more than 35-40 polyglutamine repeats seem to be the preferential target for the 1C2 antibody (Lopes-Cendes et al.). In addition, under some experimental conditions, the intensity of the abnormal band was similar to the immunoreactivity of the TBP, which is comparable to the intensity of abnormal bands observed in neurodegenerative diseases caused by CAG expansions (figure 2). Together, these observations suggest that the additional protein bands observed in the two schizophrenic patients S1 and S4 and in one of their relatives diagnosed with depression correspond to rare isoforms of a

protein with either an enlarged (upper limit of the normal size range) or an expanded (CAG repeat > upper limit the normal size range) polyglutamine tract.

Because there are slight differences in the size of the additional bands observed in S1, S4 and R2, we cannot exclude the possibility that the bands detected in the three subjects may represent different proteins. An alternative explanation to these size differences could be that they correspond to the same protein with different sizes of polyglutamine tract expansions (a short expansion will result in a smaller molecule and a weaker signal). If the latter hypothesis is true, the slightly smaller size of the novel band in R2 may reflect the fact that the schizophrenic proband inherited a larger CAG repeat leading to a more severe phenotype than his sibling with major depression. In keeping with this hypothesis, S4, who showed a slightly smaller and weaker band than S1, did not show evidence for increased psychiatric morbidity in his relatives and had a later age at onset. Further experiments are required to establish whether these proteins are the same or not.

The co-occurrence of novel protein bands in two siblings suffering respectively from schizophrenia and major depression and its absence from three non-affected siblings suggest that the corresponding protein may contribute to the increased psychiatric morbidity in this family. This is in keeping with a study showing that expanded CAG repeats are not specific to schizophrenia but are also observed in patients with affective disorders [11]. This is also consistent with epidemiological data indicating that affective disorders and schizophrenia may share some genetic susceptibility factors [22]. The absence of the abnormal protein band in one sibling who suffered from postpartum depression may result from etiological heterogeneity, which is believed to be common in such prevalent conditions, or a smaller expansion not detected with the 1C2 antibody, a frequent occurrence even in the spinocerebellar ataxias.

Although abnormal bands are expected to be more frequent in severely affected and neuroleptic nonresponsive patients, none of these tested

positive for a novel band. It is difficult to speculate on the significance of this observation given the small size of the samples and the number of detected abnormal bands.

In conclusion, only 2 out of 57 unrelated schizophrenic patients showed an aberrant band detected by the 1C2 antibody. This observation indicates that the putative expanded polyglutamine containing protein may contribute to the development of schizophrenia only in a small proportion of patients. This is consistent with the fact that schizophrenia is a complex disease where genes with a major effect are thought to be very rare.

The identification of a correlation between the presence of polyglutamine containing proteins and intranuclear inclusions in schizophrenia patients strongly suggest that these proteins are responsible for the diseased phenotype. Of importance, such a correlation had never been observed in a neuropsychiatric disease. All the reports showing such inclusion bodies had been limited to neurodegenerative diseases.

Further investigation, particularly, the identification and analysis of the gene coding for the protein corresponding to the abnormal bands and the screening of a larger group of schizophrenic patients and carefully matched controls may help to clarify their role in schizophrenia and possibly affective disorders.

CONCLUSION

A novel protein band was detected using the 1C2 antibody, which preferentially detects proteins containing expanded CAG tracts (ECAGT), in 2 patients with schizophrenia and one with major depression. Furthermore, lymphoblasts from these three patients showing the 50kDa band were analyzed for the presence of intranuclear inclusions (INI) using an anti-ubiquitin antibody. INI identical to those found in other diseases with ECAGT are found in these patients cells, but not in 20 controls. This novel finding strongly supports the hypothesis that the 50 kDa band represents a pathogenic ECAGT. The method

is a very effective way to test patients with schizophrenia for this type of mutation and can be expanded to testing patients with many other diseases.

Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims.

REFERENCES

1. Kendler et al., *Arch. Gen. Psychiatry* 1993 ;50:527-540
2. Cannon et al., *Arch. Gen. Psychiatry.* 1998; 55:67-74.
3. Kety et al., *Arch. Gen. Psychiatry.* 1994; 51:442-455.
4. Risch N., *Genet. Epidemiol.* 1990; 7:3-16; discussion 17-45.
5. Bassett et al., *Am. J. Hum. Genet.* 1997; 60:630-637.
6. Yaw et al., *Psychiatric genet.* 1996; 6:7-11.
7. Gorwood et al., *Am. J. of Psychiatry.* 1996; 153:1173-1177.
8. Chotai et al., *Psychiatric genet.* 1995; 5:181-186.
9. Petronis et al., *Am. J. of Psychiatry.* 1995; 152:164-172.
10. O'Donovan et al., *Psychological Med.* 1996; 26:1-6.
11. O'Donovan et al., *Psychological Med.* 1996; 26:1145-1153.
12. Morris et al., *Hum. Mol. Genet.* 1995; 4:1957-1961.
13. O'Donovan et al., *Nature Genet.* 1995; 10:380-381.
14. Paulson et al., *Neuron* 1997; 19:333-344.
15. Trottier et al., *Nature* 1995; 378:403-406.
16. Stevanin et al., *Hum. Mol. Genet.* 1996; 5:1887-1892.
17. Imbert et al., *Nature Genet.* 1996; 14:285-291.
18. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Diseases, Fourth Edition. 1994.
19. Paulson et al., *Annu. Rev. Neurosci.* 1996; 19:79-107.
20. Schurhoff et al., *Psychiatry Res.* 1997; 72:141-144.
21. Jones et al., *Mol. Psychiatry* 1998; 2:478-482.
22. Maier et al., *Arch. Gen. Psychiatry* 1993; 50:871-883.
23. Roizin et al., *Adv. Neurol.* 23:95-122 (1979).
24. Jackson et al., *Neuropat. Applied Neurobiol.* 21:18-26, (1995).
25. DiFiglia et al., *Science* 277:1990-1993 (1997).
26. Holmberg et al., *Human Molecular Genetics* 913-918 (1998).
27. Davies et al., *Cell* 90:537-548 (1997).
28. Ordway et al., *Cell* 91:753-763 (1997).

29. Skinner et al., *Nature* 389:971-974 (1997).

30. Davies et al., *Lancet* 351:131-133 (1998).

WHAT IS CLAIMED IS:

1. A method of diagnosing the presence or predisposition to develop a neuropsychiatric disorder comprising:

a) taking a sample from a patient;

b) determining the presence of polyglutamine containing proteins in said sample; and

c) diagnosing the presence or predisposition to develop a neuropsychiatric disorder, wherein a presence of polyglutamine containing proteins in said sample as compared to a sample from a patient without said neuropsychiatric disorder is indicative of the presence or predisposition to develop said neuropsychiatric disorder.

2. The method of claim 1, wherein said neuropsychiatric disorder is one of schizophrenia and major depression.

3. The method of claim 1 or 2, wherein said determining of the presence of polyglutamine containing proteins is carried out with a ligand specific to a polyglutamine array of said proteins.

4. The method of claim 3, wherein said ligand is an antibody.

5. The method of claim 4, wherein said antibody is monoclonal antibody 1C2.

6. A diagnostic kit comprising:

a) a first container means containing a ligand specific to a polyglutamine array in a protein; and

b) a second container means containing a conjugate comprising a binding partner of said ligand and a detectable label.

7. The diagnostic kit of claim 6, wherein said ligand is monoclonal antibody 1C2.

8. A method to identify modulators of the biological activity of polyglutamine-containing proteins associated with a neuropsychiatric disease, said method comprising:

a) an incubation of cells from a patient harboring said polyglutamine-containing proteins with a molecule; and

b) assessing the effect of said molecule on said biological activity of said polyglutamine containing proteins, wherein a difference in said biological activity in the presence of said molecule as compared to in its absence is indicative of a modulating activity of said molecule.

9. A method of treatment of a neuropsychiatric disorder comprising an administration of a polyglutamine sequence-specific ligand.

10. The method of claim 9, wherein said neuropsychiatric disorder is one of schizophrenia and major depression.

11. The method of claim 9 or 10, wherein said ligand is an antibody.

12. The method of claim 11, wherein said antibody is monoclonal antibody 1C2.

13. The method of claim 4, wherein said ligand is an anti-ubiquitin antibody or an antibody which is specific for the detection of INI.

14. A method of identification of the cell type implicated in the development of a neuropsychiatric disorder comprising an identification of the cell type containing INI in a patient suffering from a neuropsychiatric disease as compared to its absence in said cell type in a control patient.

15. Any of the above claim, wherein an animal model is used.

16. A method of diagnosing the presence or predisposition to develop a neuropsychiatric disorder comprising:

a) taking a sample from a patient;

b) determining the presence of an amino acid stretch in a protein encoded by a CAG/CTG repeat in said sample; and

c) diagnosing the presence or predisposition to develop a neuropsychiatric disorder, wherein a presence of said amino acid stretch encoded by said CAG/CTG repeat in said sample, as compared to a sample from a patient without said neuropsychiatric disorder, is indicative of the presence or predisposition to develop said neuropsychiatric disorder.

17. A diagnostic kit comprising:

a) a first container means containing a ligand specific to an amino acid stretch in a protein encoded by a CAG repeat; and

b) a second container means containing a conjugate comprising a binding partner of said ligand and a detectable label.

18. A method to identify modulators of the biological activity of a protein containing an amino acid stretch encoded by a CAG/CTG repeat associated with a neuropsychiatric disease, said method comprising:

a) an incubation of cells from a patient harboring said protein comprising said amino acid stretch with a molecule; and

b) assessing the effect of said molecule on said biological activity of said amino acid stretch encoded by said CAG/CTG repeat in said protein, wherein a difference in said biological activity in the presence of said molecule as compared to in its absence is indicative of a modulating activity of said molecule.

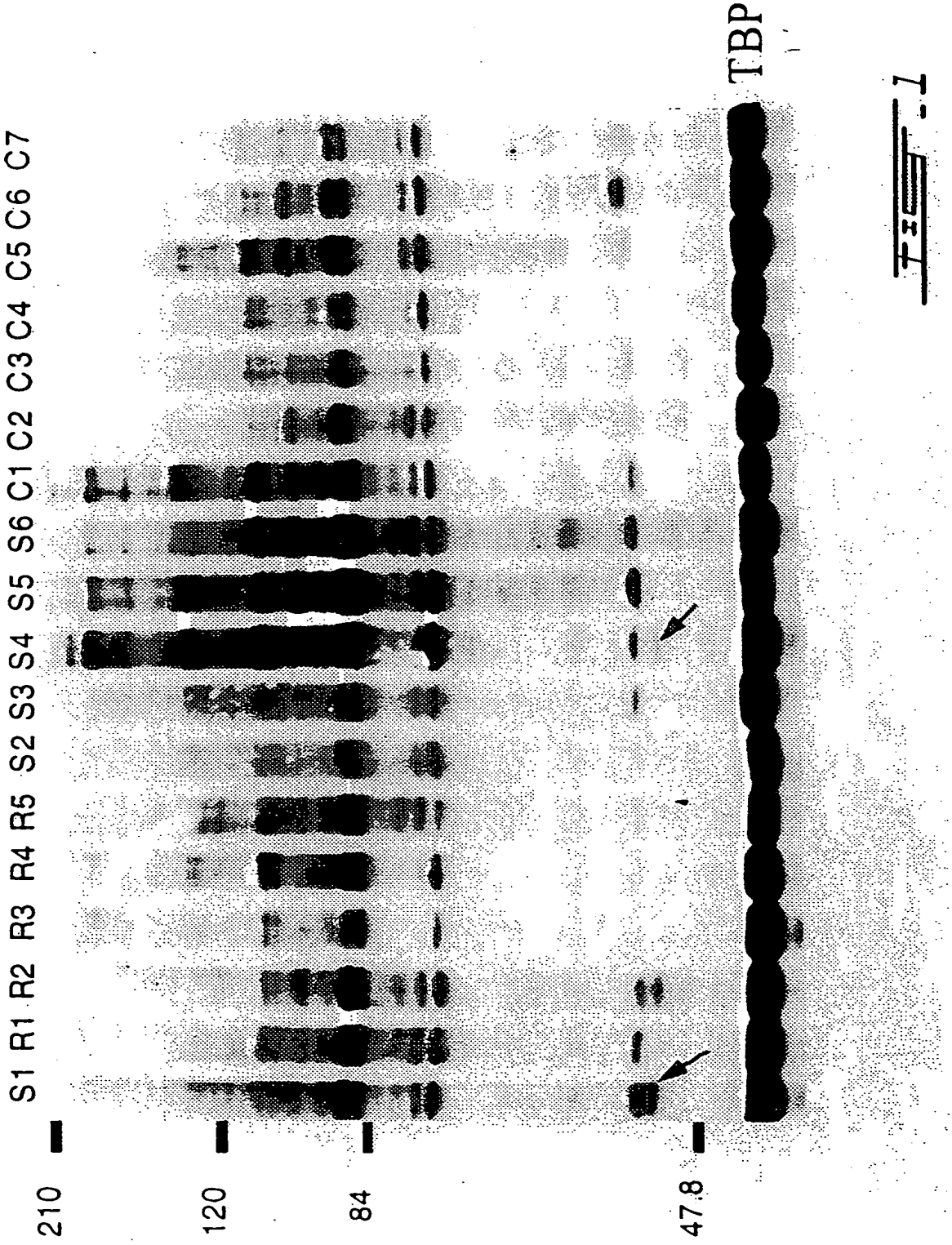

S1 R1 R2 R3 R4 R5 S2 S3 S4 S5 S6 C1 C2 C3 C4 C5 C6 C7
TBP
210
120
84
47.8

Fig. 1

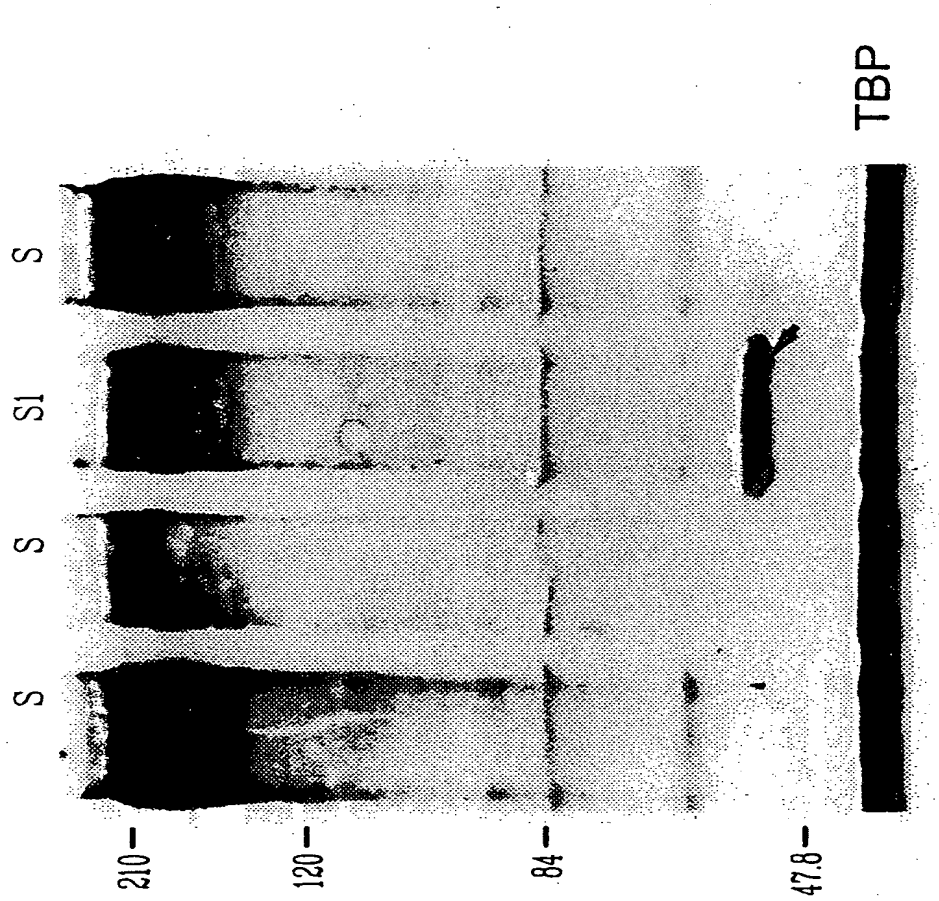


Fig-2

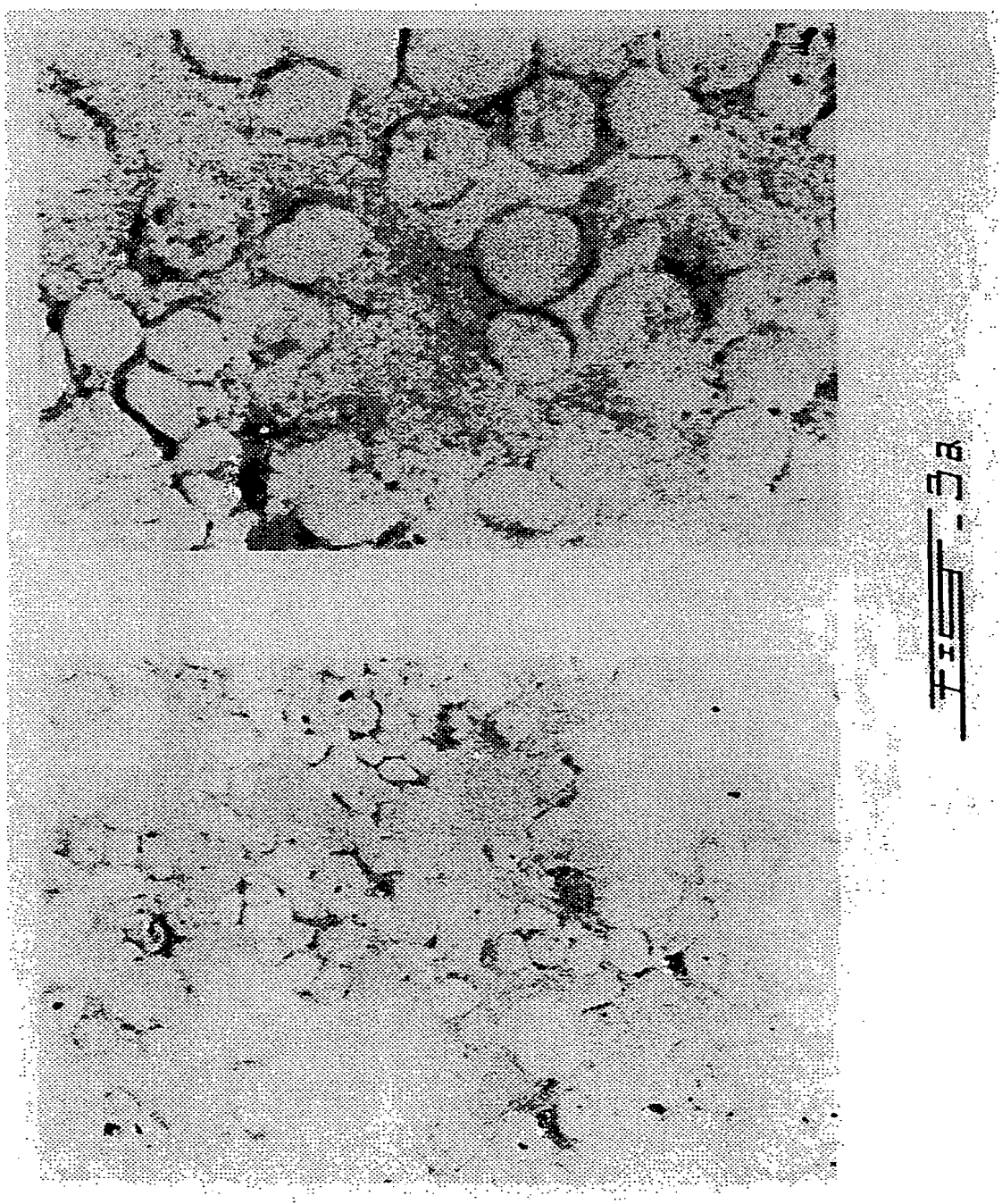


FIG. 3a

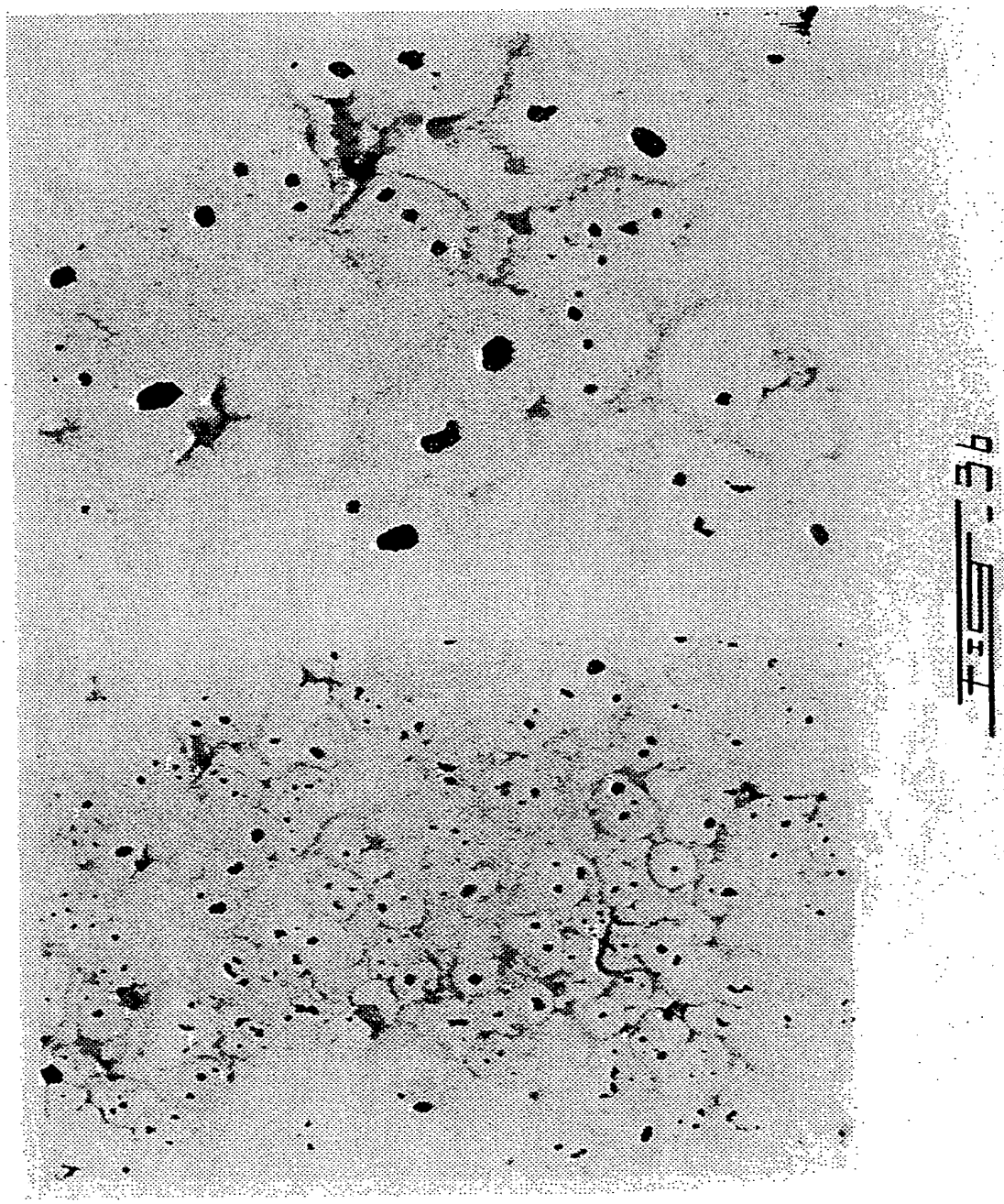


Fig. 3b

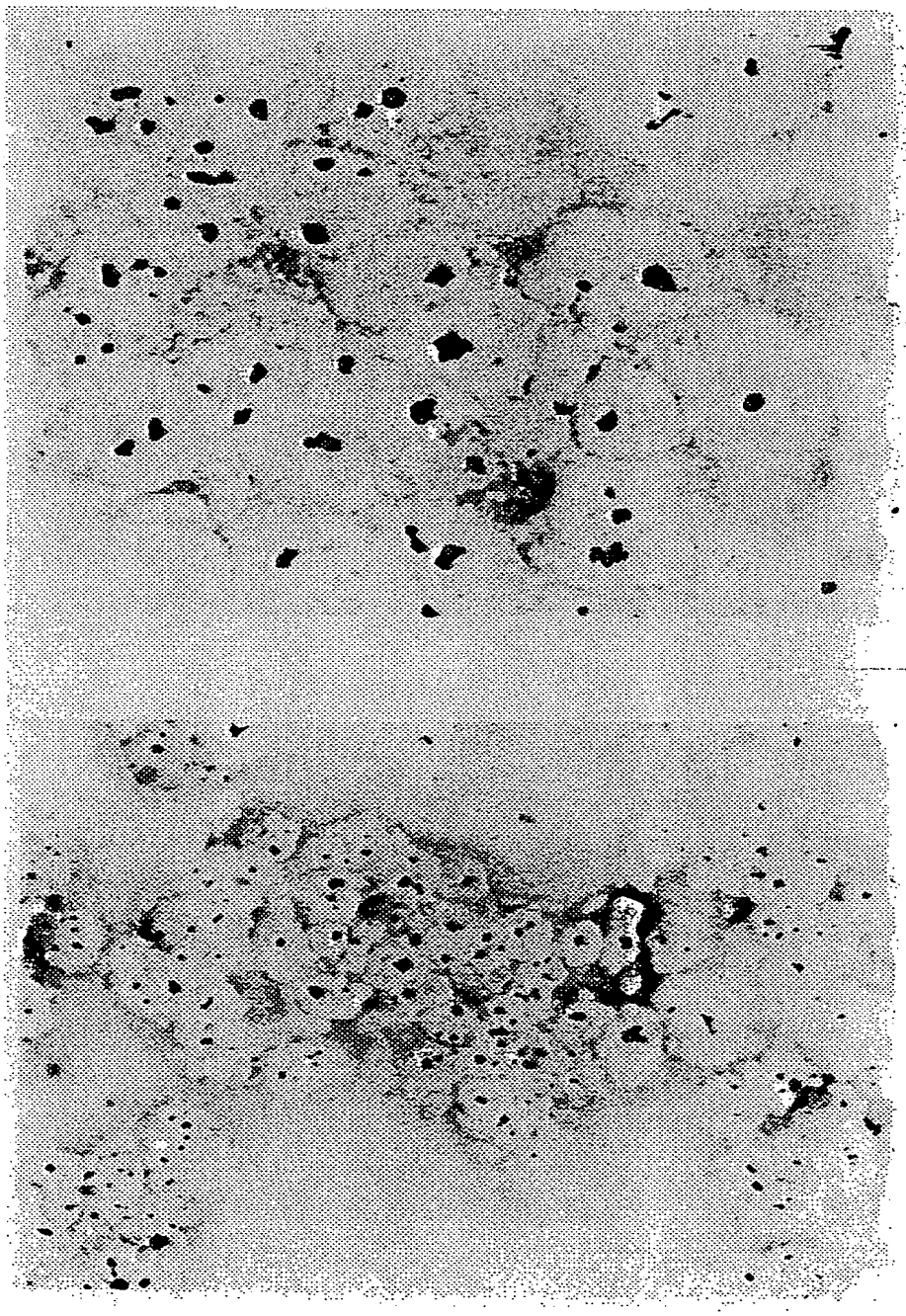


FIG. 3C

INTERNATIONAL SEARCH REPORT

Inte Application No
PCT/CA 99/01038

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/68 G01N33/577 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 723 301 A (STRITTMATTER WARREN J ET AL) 3 March 1998 (1998-03-03) the whole document	1,8,16, 18
P,X	FR 2 764 611 A (FONDATION JEAN DAUSSET CEPH) 18 December 1998 (1998-12-18) abstract	1,16,18
X	FR 2 741 088 A (CENTRE NAT RECH SCIENT) 16 May 1997 (1997-05-16) abstract	1,4,5

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 February 2000	21/02/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Moreno, C

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category '	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US JOOBER, R. ET AL: "Polyglutamine -containing proteins in schizophrenia." retrieved from STN XP002129959 abstract & MOLECULAR PSYCHIATRY, (JAN., 1999) VOL. 4, NO. 1, PP. 53-57. ,	1-18
A	DATABASE CHEMABS 'Online! CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US KOSHY, BEENA T. ET AL: "The CAG/ polyglutamine tract diseases: gene products and molecular pathogenesis" retrieved from STN Database accession no. 127:104848 XP002129960 abstract & BRAIN PATHOL. (1997), 7(3), 927-942 ,	1,6,8, 14,16-18
A	DATABASE MEDLINE 'Online! US NATIONAL LIBRARY OF MEDICINE (NLM), BETHESDA, MD, US MARGOLIS R L ET AL: "cDNAs with long CAG trinucleotide repeats from human brain." retrieved from STN Database accession no. 97369492 XP002129961 abstract & HUMAN GENETICS, (1997 JUL) 100 (1) 114-22. ,	1,6,8, 14,16-18
A	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US NERI, CHRISTIAN (1) ET AL: "Triplet repeats, neurodegenerative and neuropsychiatric diseases: Mechanisms and candidate genes." retrieved from STN XP002129962 abstract & M-S (MEDECINE SCIENCES), (1996) VOL. 12, NO. 12, PP. 1361-1369. ,	1,6,8, 14,16-18

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte[redacted] Application No
PCT/CA 99/01038

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5723301 A	03-03-1998	NONE	
FR 2764611 A	18-12-1998	WO 9856950 A	17-12-1998
FR 2741088 A	16-05-1997	WO 9717445 A	15-05-1997